7.2 Reellwertige Funktionen mehrerer Veränderlicher

2.1 Grundlagen. *Sei $D \subseteq \mathbb{R}^n$ und $f : D \to \mathbb{R}$ eine reellwertige Funktion. Dieser Funktion ist ein* ***Graph***

$$\Gamma_f := \{(\mathbf{x}, f(\mathbf{x})); \mathbf{x} \in D\} \subseteq \mathbb{R}^{n+1}$$

zugeordnet. Zur Veranschaulichung von f betrachtet man ***Niveaumengen****, d.h. für $c \in \mathbb{R}$*

$$\mathbf{N}_c := \{\mathbf{x} \in D; f(\mathbf{x}) = c\} \subseteq \mathbb{R}^n ,$$

oder Schnitte mit zu Koordinatenachsen parallelen Geraden, d.h. man betrachtet die „partielle“ Funktion von f, die gegeben ist durch

$$x_i \to f(a_1, \ldots , a_{i-1}, x_i, a_{i+1}, \ldots , a_n) ,$$

mit $\mathbf{a} = (a_1, \ldots , a_n)^T \in D$.

2.2 Definition. *Für zwei Punkte $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ wird ihr* ***Abstand*** *definiert durch*

$$\|\mathbf{x} - \mathbf{y}\| := \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}} .$$

Sei $\mathbf{a} \in \mathbb{R}^n$ ein beliebiger aber fester Punkt und $r > 0$, dann heißt die Menge

$$B_r(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n; \|\mathbf{x} - \mathbf{a}\| < r\}$$

*r-**Umgebung** von $\mathbf{a}$.*

2.3 Definition. *Sei D eine Teilmenge des $\mathbb{R}^n$.*

a) *Ein Punkt $\mathbf{a} \in D$ heißt* ***innerer Punkt*** *von D, wenn es eine r-Umgebung $B_r(\mathbf{a})$ von $\mathbf{a}$ gibt, die ganz in D liegt, d.h. $B_r(\mathbf{a}) \subseteq D$.*

b) *D heißt* ***offen****, wenn jeder Punkt von D ein innerer Punkt ist.*

c) *Ein Punkt $\mathbf{b} \in \mathbb{R}^n$ heißt* ***Randpunkt*** *von D, wenn jede r-Umgebung von $\mathbf{b}$ sowohl mindestens einen Punkt aus D als auch einen nicht zu D gehörenden Punkt enthält. Die Menge aller Randpunkte heißt* ***Rand*** *von D und wird mit ∂D bezeichnet.*

d) Eine Menge ist abgeschloßen, wenn sie alle ihre Randpunkte enthält.

Beispiele:

- Die Kreisscheibe $B_r(\mathbf{a}) = \{(x,y)^T \in \mathbb{R}^2, (x-a_1)^2 + (y-a_2)^2 < r^2\}$ ist offen. Der Rand ist $\{(x,y)^T; (x-a_1)^2 + (y-a_2)^2 = r^2\}$.
- $C_r(\mathbf{a}) = \{(x,y)^T \in \mathbb{R}^2; (x-a_1)^2 + (y-a_2)^2 \leq r^2\}$ ist abgeschloßen.
- $D = \mathbf{B}_r(\mathbf{a}) \cap \{(x,y)^T \in \mathbb{R}^2; x \geq 0\}$ weder offen noch abgeschloßen.

2.4 Definition. *Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{a} \in D \cup \partial D$.*

a) f hat in $\mathbf{a}$ den Grenzwert $c \in \mathbb{R}$, in Zeichen

$$\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = c$$

wenn es zu jedem $\varepsilon > 0$ eine r-Umgebung $B_r(\mathbf{a})$ gibt, so daß

$$|f(\mathbf{x}) - c| \leq \varepsilon\,.$$

für alle $\mathbf{x} \in D \cap B_r(\mathbf{a})$ gilt.

*b) f heißt in $\mathbf{a} \in D$ **stetig**, wenn $\lim_{\mathbf{x} \to \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ gilt.*

*c) f heißt auf D **stetig**, wenn f in allen $\mathbf{a} \in D$ stetig ist.*

- *Analog zum Fall $n = 1$ gelten für Grenzwerte und stetige Funktionen die üblichen Rechenregeln, d.h. Summe, Produkt und Quotient stetiger Funktionen sind stetig.*
- Die Projektion $p_i(\mathbf{x}) = x_i$ ist stetig und somit sind alle Polynome

  $$p(\mathbf{x}) = \sum_{1 \leq k_i \leq m} a_{k_1,\dots,k_n} x_1^{k_1} \cdots x_n^{k_n}$$

  stetig. Insbesondere sind lineare Abbildungen

  $$\ell(\mathbf{x}) = \mathbf{a}^T \mathbf{x} = \sum_{i=1}^{n} a_i x_i$$

  stetig. Die rationale Funktion $r(\mathbf{x}) = \frac{p(\mathbf{x})}{q(\mathbf{x})}$ ist stetig, falls $q(\mathbf{x}) \neq 0$ für alle $\mathbf{x} \in D$.

- Achtung! Aus der Stetigkeit der partiellen Funktionen folgt nicht die Stetigkeit von f.

2.5 Definition.

a) *Eine Menge $D \subseteq \mathbb{R}^n$ heißt beschränkt, wenn es eine Konstante $K > 0$ gibt mit*

$$\|\mathbf{x}\| < K \qquad \textit{für alle } \mathbf{x} \in D\,.$$

b) *Die abgeschlossenen und beschränkten Mengen des $\mathbb{R}^n$ nennt man **kompakt**.*

2.6 Satz. *Jede auf einer kompakten Menge $D \subseteq \mathbb{R}^n$ stetige Funktion nimmt auf D ein Minimum und ein Maximum an, d.h. es gibt $\mathbf{a}, \mathbf{b} \in D$ mit*

$$f(\mathbf{a}) \le f(\mathbf{x}) \le f(\mathbf{b}) \qquad \textit{für alle } \mathbf{x} \in D\,.$$

2.7 Satz. *Jede auf einer kompakten Menge $D \subseteq \mathbb{R}^n$ stetige Funktion ist **gleichmäßig stetig**, d.h. zu $\varepsilon > 0$ gibt es ein $\delta > 0$, so daß für alle $\mathbf{x}, \mathbf{y} \in D$ gilt:*

$$\|\mathbf{x} - \mathbf{y}\| < \delta \quad \Rightarrow \quad |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon\,.$$

- Die Sätze 2.6 und 2.7 sind im Allgemeinen falsch, wenn D nicht kompakt ist.

- Sei $f : D \to \mathbb{R}$ stetig, aber D nicht kompakt. Falls für alle $\mathbf{x} \in \partial D$ die Funktion f in $\mathbf{x}$ einen Grenzwert hat, kann man f stetig auf $D \cup \partial D$ fortsetzen durch

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \mathbf{x} \in D, \\ \lim\limits_{\substack{\mathbf{y} \to \mathbf{x} \\ \mathbf{y} \in D}} f(\mathbf{y}) & \mathbf{x} \in \partial D. \end{cases}$$

  Dann ist $\tilde{f}$ stetig auf der kompakten Menge $D \cup \partial D$ und Satz 2.6 und Satz 2.7 gelten.

- Sei $f : D \to \mathbb{R}$ stetig und sei $f(\mathbf{a}) > 0$ für $\mathbf{a} \in D$. Dann gibt es ein $r > 0$ so daß

$$f(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in D \cap B_r(\mathbf{a})\,.$$

2.8 Definition. *Sei $D \subseteq \mathbb{R}^n$ offen, $f : D \to \mathbb{R}$ eine Funktion und $\mathbf{x} \in D$. Existiert der Grenzwert*

$$\lim_{t \to 0} \frac{1}{t} \big(f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n) \big)$$

so wird er als die ***partielle Ableitung von*** *f* ***nach*** *x_i an der Stelle $\mathbf{x}$ genannt und als $\frac{\partial f}{\partial x_i}(\mathbf{x})$ bezeichnet.*

- *Die partielle Ableitung ist die Ableitung der „partiellen" Funktionen*

  $$x_i \to f(x_1, \dots, x_i, \dots, x_n),$$

  wobei alle Werte $x_j, j \neq i$ konstant gehalten werden.

- *Eine andere Bezeichnung ist f_{x_i}.*

- ***Höhere partielle Ableitungen*** *werden analog definiert und z.B. mit $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}$ bezeichnet.*

- *f heißt* ***partiell differenzierbar*** *bzw.* ***stetig partiell differenzierbar*** *wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ existieren bzw. stetig sind.*

- *Analoge Definition für höhere Ableitungen.*

2.9 Definition. *Der* ***Gradient*** *einer Funktion f ist der Vektor bestehend aus den partiellen Ableitungen, d.h.*

$$\operatorname{grad} f(\mathbf{x}) := \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)^T \in \mathbb{R}^n .$$

- *Eine andere Bezeichnung ist: $\nabla f(\mathbf{x})$.*

- *Sei $D \subseteq \mathbb{R}^n$ offen und $k \in \mathbb{N}_0$. Dann bezeichnen wir*

  $$C^k(D, \mathbb{R}) := \{ f : D \to \mathbb{R};\, f \ k\text{-mal stetig partiell differenzierbar} \}$$

- *Im Allgemeinen gilt: $\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$.*

2.10 Satz. *Für jede C^2-Funktion $f : D \to \mathbb{R}, D \subseteq \mathbb{R}^n$ offen gilt:*

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right), \qquad 1 \leq i, j \leq n .$$

- Der Satz gilt analog für k-te gemischte partielle Ableitungen, falls $f \in C^k(D, \mathbb{R})$.

2.11 Definition. *Für zwei Funktionen $f, g : D \subseteq \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{x}_0 \in D$, $k \in \mathbb{N}_0$ schreibt man*

$$f(\mathbf{x}) = g(\mathbf{x}) + o(\|\mathbf{x} - \mathbf{x}_0\|^k) \qquad \text{für } \mathbf{x} \to \mathbf{x}_0$$

falls

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - g(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^k} = 0\,.$$

- *Sei $f : \mathbb{R} \to \mathbb{R}$ differenzierbar an der Stelle x_0. Nach Kapitel 3 1.6 ist die beste lineare Approximation von f an der Stelle x_0 gegeben durch $g(x) = f(x_0) + f'(x_0)(x - x_0)$, d.h.*

  $$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|)\,.$$

  Für $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ reicht partielle Differenzierbarkeit für das Analogon obiger Formel nicht.

2.12 Definition. *Sei $D \subseteq \mathbb{R}^n$ offen. Eine Funktion $f : D \to \mathbb{R}$ heißt in $\mathbf{x}_0 \in D$* ***total differenzierbar****, wenn es einen Vektor $\mathbf{a} \in \mathbb{R}^n$ mit*

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|), \tag{2.13}$$

für $\mathbf{x}$ nahe $\mathbf{x}_0$, gibt.

2.14 Satz. *Ist f in $\mathbf{x}_0 \in D$ total differenzierbar mit $f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|)$, dann gilt:*

a) *f ist in $\mathbf{x}_0$ stetig,*

b) $\lim_{t \to 0} \frac{1}{t} (f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)) = \mathbf{a} \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq 0,$

c) *f ist partiell differenzierbar und $\mathbf{a}$ ist eindeutig bestimmt als $\mathbf{a} = \operatorname{grad} f(\mathbf{x}_0)$.*

- Wenn f total differenzierbar ist, so ist

  $$f(\mathbf{x}) = f(\mathbf{x}_0) + \operatorname{grad} f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|) \tag{2.15}$$

  die beste lineare Approximation von f nahe $\mathbf{x}_0$.

2.16 Satz. *Jede C^1-Funktion $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, mit D offen, ist auf D total differenzierbar, d.h. jede stetig partiell differenzierbare Funktion ist total differenzierbar.*

2.17 Newton - Verfahren. *Zur näherungsweisen Lösung eines nichtlinearen Gleichungssystems*

$$f(x,y) = 0\,, \qquad g(x,y) = 0\,, \tag{*}$$

für $f, g \in C^1$ ersetzt man (*) durch die lineare Approximation

$$\begin{aligned} f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) &= 0\,,\\ g(x_0,y_0) + g_x(x_0,y_0)(x-x_0) + g_y(x_0,y_0)(y-y_0) &= 0\,. \end{aligned}$$

Dieses kann man als

$$\mathbf{A}\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}$$

schreiben, wobei

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} f_x(x_0,y_0) & f_y(x_0,y_0) \\ g_x(x_0,y_0) & g_y(x_0,y_0) \end{pmatrix}\,,\\ \mathbf{g} &= -\begin{pmatrix} f(x_0,y_0) \\ g(x_0,y_0) \end{pmatrix} + \mathbf{A}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}\,. \end{aligned}$$

Falls $\det \mathbf{A} \neq 0$ *erhält man eine verbesserte Lösung zum Startwert $(x_0, y_0)^T$ durch*

$$\begin{aligned} x_1 &= x_0 + \frac{1}{\det \mathbf{A}} \det\begin{pmatrix} -f(x_0,y_0) & f_y(x_0,y_0) \\ -g(x_0,y_0) & g_y(x_0,y_0) \end{pmatrix}\,,\\ y_1 &= y_0 + \frac{1}{\det \mathbf{A}} \det\begin{pmatrix} f_x(x_0,y_0) & -f(x_0,y_0) \\ g_x(x_0,y_0) & -g(x_0,y_0) \end{pmatrix}\,. \end{aligned}$$

Dieses Vorgehen kann wiederholt werden und liefert oft gute Ergebnisse bei entsprechender Wahl des Startpunktes.

2.18 Definition. *Sei $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ und $\mathbf{x}_0 \in D$ und $\mathbf{v} \in \mathbb{R}^n$ ein Einheitsvektor, d.h. $\|\mathbf{v}\| = 1$. Falls der Grenzwert*

$$\lim_{t\to 0} \frac{1}{t}\left(f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)\right)$$

*existiert wird er mit $\partial_{\mathbf{v}} f(\mathbf{x}_0)$ oder $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ bezeichnet und heißt **Richtungsableitung** von f an der Stelle $\mathbf{x}_0$ in Richtung $\mathbf{v}$.*

- *Dies ist eine Verallgemeinerung von partiellen Ableitungen welche der Wahl* $\mathbf{v} = \mathbf{e}_i$ *entspricht.*
- $\partial_{\mathbf{v}} f(\mathbf{x}_0)$ *beschreibt das Verhalten von* f *längs der Geraden* $\mathbf{x}_0 + t\mathbf{v}$.

2.19 Satz. *Für jede auf der offenen Menge* $D \subseteq \mathbb{R}^n$ *total differenzierbare Funktion* f *und alle* $\mathbf{v} \in \mathbb{R}^n, \|\mathbf{v}\| = 1$ *gilt:*

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \text{grad } f(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{i=1}^{n} f_{x_i}(\mathbf{x}_0) v_i$$

• Wir hatten bereits gesehen, daß sich Funktionen längs verschiedener Kurven verschieden verhalten können (siehe Beispiel stetiger „partieller" Funktionen). Deshalb ist es sinnvoll auch Ableitungen längs von Kurven und nicht nur längs von Geraden (Richtungsableitungen) zu betrachten.

2.20 Satz (Kettenregel). *Für jede* C^1*-Funktion* $f : D \to \mathbb{R}, D \subseteq \mathbb{R}^n$ *offen, und für jedes Kurvenstück* $\mathbf{x} : [a, b] \to D$ *gilt:*

$$\frac{d}{dt} f(\mathbf{x}(t)) = \text{grad } f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \,. \tag{2.21}$$